

Setting Standards: Information Accumulation in Development*

Daron Acemoglu

Massachusetts Institute of Technology and CEPR
and

Fabrizio Zilibotti

Universitat Pompeu Fabra and CEPR

March, 1997

Abstract

We propose a model in which economic relations and institutions in advanced and less developed economies differ as these societies have access to different amounts of *information*. This lack of information makes it hard to give the right *incentives* to managers and entrepreneurs. We argue that differences in the amount of information arise because of the differences in the scale of activities in rich and poor economies; namely, there is too little *repetition* of similar activities in poor economies, thus insufficient information to *set the appropriate standards* for firm performance. Our model predicts a number of institutional and structural transformations as the economy accumulates capital and information.

Keywords: Information, Development, Agency Costs, Incentives, Relative Performance Evaluation, Risk Sharing, Sectorial Transformations.

JEL Classification: D82, M13, O13, O14, O40.

We thank Piero Gottardi, Dilip Mookherjee, Jaume Ventura and seminar participants at Boston University, IGIER-Bocconi for useful comments. Zilibotti acknowledges the kind hospitality and financial support of the European Forum of the EUI.

1 Introduction

The traditional approach views economic development as a set of interrelated changes ranging from the structure of production to social institutions. In this paper we propose a theory of development based on the evolution of principal-agent relations which emphasizes some of the same themes. Even though we are far from providing a unifying theory that can do justice to such a complex phenomenon, our model will generate a host of structural changes in the process of growth. In particular, in our economy, efficiency and productivity will increase, the extent of risk sharing will change, agents will use more sophisticated production techniques and more division of labor, the sectorial composition of output and the degree of specialization in the economy will evolve, and some new financial institutions will emerge while others disappear. Naturally, our analysis and therefore these results will be extremely stylized, but we will argue, they give some of the flavor of the complex process that is development.

Our main argument can be broken into the following steps:

1. Delegation of tasks, employment relations and entrepreneurial activities are important for wealth creation. All of these give a first-order role to principal-agent relations in the organization of production (see for instance Mokyr, 1991, North, 1990, Pollard, 1965, Stiglitz, 1987).
2. Principal-agent relations fail when information is scarce and thrive when information is abundant. In particular, in the absence of adequate information, it is excessively costly to give the right incentives to workers, managers and entrepreneurs (e.g. Holmstrom, 1979, for a formal analysis).
3. Societies *accumulate information* about a certain activity by repeating it, in other words, by allocating some of their scarce resources to this line of business. When a certain activity is repeated many times, the society will have developed a *standard* for it. This standard can be used by the entrepreneurs to increase their efficiency, or by principals to control the entrepreneurs/managers by comparing their performance to this standard. This view is echoed by the historical study of Pollard (1965) who explains that embezzlement and inefficiency were very common among

managers at first, led to high failure rates in mining, textiles and other new industries during the seventeenth century, and were eliminated only slowly over time. In the end, professional management became part of the efficient organization of production. Regarding the importance of setting *standards*, Pollard writes: “*as ever there was considerable difference between the pioneers, ploughing a lone furrow, and those in groups large enough to develop a body of knowledge...*” (p. 122). A special case of developing a standard is the use of relative performance evaluation which we will use as our main example to illustrate how information improves principal-agent relations.

4. In a less developed economy, capital (broadly construed) is scarce, and therefore cannot be allocated in great quantity to every single activity. The result is lack of *repetition* in many areas and thus informational scarcity. As the economy accumulates capital, it will start devoting more resources to a range of activities, and the accumulation of information in each activity will improve the principal-agent relations and productivity.

These steps bring us to our first main conclusion: prosperity implies more information which implies greater efficiency, and, in turn again, prosperity. This is a process of growth that would be called a “virtuous circle” by Singer (1949) or “circular cumulative causation” by Myrdal (1956). The mechanism through the accumulation of information and change of incentives is an important, and to date unexplored, alternative to the existing formalization of these ideas (e.g. Murphy, Shleifer and Vishny, 1989, or Matsuyama, 1994). What distinguishes our model is not only that we are offering an alternative microfoundation for “sectorial externalities”, but the implications. The improvements principal-agent relations due to information accumulation lead to a number of the structural transformations which are not predicted by existing models.

Before discussing the implications of our approach, we outline our formal model. Our economy has many sectors (islands). Production in each sector takes place within firms run by entrepreneurs (managers) using capital and labor. The output of each firm depends on managerial effort, an idiosyncratic shock and a sector-specific shock. The effort choice of the entrepreneur is her private information, introducing the principal-agent problem in our econ-

omy. She can be induced to exert effort if her compensation depends on her performance. Since all agents are risk-averse, this is costly. As the capital stock increases, more capital will be allocated to different activities, enabling the parallel employment of more entrepreneurs in each sector. With many entrepreneurs in a sector, average performance can be used as an adequate standard and will filter out sector-specific shocks ensuring better incentives. As a result, the level of entrepreneurial effort and the level of productivity will be increasing in the aggregate capital stock of the economy. Since productivity is endogenously lower at earlier stages of development, growth will be slower relative to a pure neoclassical version of our model.

Now the implications:

- a) We find that as the economy develops and more information is accumulated, there are two opposing forces impacting on the extent of *risk sharing*. On the one hand, more information enables greater risk sharing at a given level of effort. On the other hand, with more information entrepreneurs are induced to exert more effort which reduces risk sharing. As a result, risk sharing may decline with development or follow an inverse U shape pattern. This result may explain the interesting findings of Townsend (1995a,b). In his studies of Asian villages he finds the degree of risk sharing is lower in richer villages; “*as if consumption insurance, whether indigenous or otherwise, deteriorates with growth*” (Townsend, 1995b, p. 95). We know of no other explanation for this interesting finding.
- b) Our model predicts that a less developed economy may be highly specialized in order to economize on agency costs by investing only in a few sectors. The structure of production will then become more balanced as the economy develops. In particular, when the importance of effort and the nature of uncertainty vary across activities, relative production and productivity in the sectors with more agency problems will increase with development. It is a well-known pattern that the share of agriculture is strongly correlated with per capita income, and this is often explained by some sectorial “increasing returns” in manufacturing (Matsuyama, 1991, see Backus, Kehoe and Kehoe, 1992, for some evidence). Our model offers a microfoundation for this sectorial increasing returns in manufacturing whereby, even though technologically unbiased, growth will favor

some sectors at the expense of others.

- c) Since Adam Smith, division of labor is seen as an important engine of growth. If the process of production can be separated into smaller or more specialized tasks and delegated to different agents, efficiency can be improved. And yet, with division of labor there is a need to provide the right incentives to the agents working in different parts of the production process. Our model shows that as information is accumulated and agency costs decline, production with division of labor (delegation) will become more attractive. The more general prediction is that a developed society should use more “sophisticated” production techniques than less developed economies.
- d) Economic development is typically accompanied by urbanization. Bairoch (1988, pp. 246-248), for instance, describes how England (and many other countries) got transformed from an essentially rural society to an urban one. Our model enables us to think of another dimension of this transformation. The close-knit structure of villages provides “efficient” direct *monitoring* of agents, which is certainly related to Townsend’s and other researchers’ findings that simple village institutions achieve a high degree of consumption risk sharing. In contrast to villages, monitoring agents in cities is much harder, hence the common notion that cities provide more privacy to individuals. Our model then predicts that at the early stages of development, the close monitoring of villages may be very valuable, but this advantage diminishes as more decentralized information is accumulated, and the right incentives can be given to entrepreneurs in cities without excessive costs. More generally, this reasoning explains a common perception, expressed succinctly by Rosenberg and Bridzell (1985): “*The move to wealth is a move towards greater possibilities of privacy and individual choice*” (p. 4).
- e) A similar reasoning to (d) also gives us a way of understanding the transformation of financial arrangements over the process of development. In a classic historical study, Goldsmith (1987) shows that premodern societies had institutions to intermediate funds, but these were quite different from what we have today. Intermediation was local and relied on heavy

monitoring. In contrast, today a large fraction of funds are intermediated by stock and bond markets and even the banks which still play an important role do little monitoring relative to village “usurers” of older times. The same pattern is observed in a cross-section when the financial institutions of low income economies are compared to their western counterparts (e.g. Besley, 1995). Again, when there is little information in the economy regarding how a certain business should be conducted, close monitoring is beneficial. As a larger scale of activity increases the amount of decentralized information, different institutions become relatively more efficient.

As the discussion suggests, our paper is related to many different strands of the literature, and without doing full justice to the insights of some of these. The growth aspect of our model is rather standard. More important for our paper is the structure we borrow from information economics. We model relative performance evaluation using the linear structure first introduced by Holmstrom and Milgrom (1987) (from now on H-M, see also the applications in H-M, 1991). Models of relative performance evaluation date back even further, for instance, Lazear and Rosen (1981), Holmstrom (1982), Green and Stokey (1983) and Shleifer (1985). The key differences between our paper and these contributions are: (a) the fact that the information is endogenous in our setting; and (b) the application of endogenous information accumulation to the process of development.

Other papers closely related to our work are Greenwood and Jovanovic (1990), Greenwood and Smith (1993), Acemoglu and Zilibotti (1996,97), Banerjee and Newman (1993,96). Greenwood and Jovanovic (1990), Greenwood and Smith (1993) and Acemoglu and Zilibotti (1996) discuss issues of financial development and the last two compare the roles of banks and stock markets. Acemoglu and Zilibotti (1997) explain some patterns of the development process by exploiting the endogenous risk-return trade-off but does not feature agency problems. Banerjee and Newman (1993) account for the patterns of occupational choice over the process of development using a model of agency. In Banerjee and Newman (1996) they explain the shift of population from villages to cities, based on the assumption that agency costs are less severe in villages than cities. The main difference between their approach, which also

brings back some of the important insights of the older literature on development economics, and ours is that they concentrate on the general equilibrium implications of wealth effects on the interest rate. In particular, when agents are poor, they have no collateral, and cannot borrow, but with development, these borrowing constraints are relaxed. Since like all previous contributions, accumulation of information is absent in their papers, their mechanism is different but complementary to ours.

The plan of our paper is as follows: section 2 describes the environment and characterizes the equilibrium in the absence of imperfect information. Section 3 is the core of the paper. It introduces imperfect information, characterizes the equilibria and determines the impact of capital accumulation on the organization of production and productivity. Section 4 discusses a number of implications of our model. Section 5 concludes and the Appendix contains all the proofs.

2 The Model Without Informational Imperfections

2.1 Technology and Preferences

We consider an economy populated by a sequence of one-period lived altruistic generations. Each generation consists of a continuum N of agents where $N \gg 1$. Throughout his life, each agent inherits and invests his (her) parent's savings, earns labor income and makes savings decisions. The utility of an agent of dynasty h born in period t is given as:

$$u(c_t^h, b_t^h, e_t^h) = -\exp \left[-\bar{\rho} \left((b_t^h)^s (c_t^h)^{1-s} - \frac{1}{2\beta} (e_t^h)^2 \right) \right] \quad (1)$$

where e_t^h is effort, c_t^h denotes consumption and b_t^h denotes the funds left for bequest. These funds are invested at the market rate of return, and the returns accrue to the offspring as bequest. Two features are worth noting. First, agents care about the amount of bequest they leave rather than about their offspring's utility (Andreoni, 1989). Second, preferences over total wealth exhibit Constant Absolute Risk Aversion (CARA).

Each agent has a career choice. He can become either a worker and earn

the wage w , or a manager (entrepreneur) and earn the managerial salary z . His total income will then be given as follows:

$$x_t^h = \begin{cases} r_t b_{t-1}^h + w_t & \text{if worker} \\ r_t b_{t-1}^h + z_t & \text{if manager} \end{cases}$$

where x denotes income, and r_t denotes the gross rate of return on savings. We will further assume that capital depreciates upon use, thus r_t will also be the net rate of return on savings. Given the separability between the saving and effort decisions, utility maximization implies that $c_t^h = (1-s)x_t^h$ and $b_t^h = s x_t^h$, thus the warm-glove bequests and Cobb-Douglas preferences ensure a constant savings rate s which will simplify the dynamics in our model. The indirect utility of agent h can then be written as:

$$U(x_t^h, e_t^h) = -\exp \left[-\rho \left(x_t^h - \frac{1}{2\beta} (e_t^h)^2 \right) \right] \quad (2)$$

where we have defined $\rho \equiv s^s(1-s)^{1-s}\bar{\rho}$ and $\beta \equiv s^s(1-s)^{1-s}\bar{\beta}$. When this will cause no confusion, we will drop superscript h .

There is a continuum 1 of islands. Every period there are N agents on each island $j \in [0, 1]$. Labor cannot be transferred between islands but capital and final output can. Therefore, each of the N agents in island j will have to work there, but they can invest their capital in any of the islands of this economy. Within the island, some of the agents will choose to be managers while some others become workers. All islands produce exactly the same good.

Production requires *one* manager, labor and capital. The amount of final good produced by firm i in island j at time t , y_{ijt} , is given by:

$$y_{ijt} = (\mu + e_{ijt} + a_{jt}) \cdot \min \left[1, l_{ijt}^\alpha k_{ijt}^{1-\alpha} \right] + \epsilon_{ijt} \quad (3)$$

where l_{ijt} is total labor hired by the firm *inclusive* of the manager, k_{ijt} is capital, μ is a constant, e_{ijt} is the effort exerted by the manager of this firm, a_{jt} is an island specific productivity shock which affects all firms on island j , and ϵ_{ijt} is an idiosyncratic shock which only affects this firm. We think of ϵ_{ijt} as capturing the importance of luck as well as the role of managerial ability ex-ante unknown to the agents. All productivity shocks are independent from each other and over time, and we have $\epsilon_{ijt} \sim N(0, \sigma^2)$ and $a_{jt} \sim N(0, \nu^2)$.

The Law of Large Numbers implies (ignoring technical details related to the continuum) that in each period $\int_0^1 a_{jt} dj = 0$ and $\int_0^1 \sum_i \epsilon_{ijt} dj = 0$.

The form of the production function captures the idea that a manager is necessary for production (“division of labor”) and has to exert some effort but also that there is only a limited amount of capital and labor that the manager can productively use (see also footnote 6 below). Though managers have to exert effort, there is no need for workers to exert positive effort (or equivalently, we could assume that workers also have to exert effort but they are perfectly monitored).

Capital is owned and supplied competitively by the agents. As we will see shortly, constant savings rate and CARA preferences over the uncertain income process will imply that wealth effects are absent, thus income distribution among agents does not matter for occupational choices and the dynamics of the aggregate capital stock. Also because capital is perfectly mobile, the distribution of wealth across islands is of no importance. Therefore, the total stock of capital, K_t , will be the unique state variable of this economy.

Finally, we assume that there exists a large set of potential intermediaries, which we refer to as *firms*. They can *freely* decide to *enter* into any of the islands of this economy and are owned at time t by generation t agents. Each active firm rents capital, labor and *one* manager for production. Both workers and managers have to be hired from a competitive spot market. We assume that each agent owns an equal share of *all* of the firms in the economy. If a firm makes positive profits, these are distributed among the owners, and if it makes a loss, the losses are also distributed among the owners. Since there is no aggregate risk in this economy, an agent who owns an equal share of *all* the firms bears no risk. Therefore, all firms will simply maximize *expected profits*. It is straightforward to see that if we started from a situation in which some agents owned a small subset of the firms and thus faced risks depending on the profits of these firms, then there would be Pareto improving trades in shares.

2.2 The Equilibrium Concept

Throughout the paper we will use a concept of equilibrium very close to the standard notion of *competitive equilibrium*, and we will model it as the equilibrium of a two-stage game in the presence of *unfettered competition* among firms

(see Townsend, 1983). In the first stage, potential firms announce in which island, if any, they will be active. If firm i announces at time t that it will be active in island j , we denote this by $i \in \mathcal{M}_{jt}$. We also let $M_{jt} \equiv \#\mathcal{M}_{jt}$ be the number of firms that have announced that they will be active in island j . In the second stage all $i \in \mathcal{M}_{jt}$ take the first stage announcements as summarized by $\mathbf{M}_t \equiv \{M_{jt}\}_{j \in [0,1]}$ and the total capital stock of the economy, K_t , as given and compete to hire workers and managers from island j and capital from the economy-wide market. We restrict each firm to hire at most one manager.¹ We also use ω_t to denote the publicly observed state of nature at time t .

A static equilibrium at time t is a set of first stage announcements summarized by \mathbf{M}_t , factor return functions $w_j(\omega_t; \mathbf{M}_t, K_t)$ and $z_j(\omega_t; \mathbf{M}_t, K_t)$ for all $j \in [0, 1]$, and $r(\omega_t; \mathbf{M}_t, K_t)$, and labor and capital demands for all $i \in \mathcal{M}_{jt}$ and $j \in [0, 1]$, $l_{ij}(\mathbf{M}_t, K_t)$ and $k_{ij}(\mathbf{M}_t, K_t)$, such that:²

1. No firm $i \in \mathcal{M}_{jt}$ for any $j \in [0, 1]$ can strictly increase its expected profits by offering a different function than w_j , z_j , and r , or by demanding different amounts of labor and capital than l_{ij} and k_{ij} .
2. No firm can change its entry decision in the first stage and strictly increase its expected profits.
3. The resulting allocation is feasible, in the sense that:

$$\int_{i \in \mathcal{M}_{jt}} l_{ij} di \leq N, \quad \forall j \in [0, 1] \quad (4)$$

$$\int_0^1 \int_{i \in \mathcal{M}_{jt}} k_{ij} di dj \leq K_t. \quad (5)$$

4. In every island $j \in [0, 1]$, given $w_j(\omega_t; \mathbf{M}_t, K_t)$ and $z_j(\omega_t; \mathbf{M}_t, K_t)$, M_{jt} agents choose to become managers.

¹We also assume that if a firm announces that it will be active in island j and does not hire a manager, it incurs a small cost ϵ . This assumption ensures that in equilibrium only firms which will be active enter.

²An equivalent definition of equilibrium can be given where all agents would be pure price-takers. This does not affect our results as long as the space of commodities is chosen appropriately, in particular, insurance contracts need to be allowed, and some incentive compatibility constraint would have to be imposed on these when there is imperfect information. Yet another equivalent model would involve managers hiring workers and capital, and writing co-insurance contracts with each other while also respecting incentive compatibility.

5. All managers choose effort to maximize their utility.

Note in particular that we have already imposed as part of the equilibrium concept that all firms in island j will pay the same (state-contingent) price for labor and managers, and all firms in the economy will pay the same (state-contingent) price for capital. This is a straightforward implication of competition among firms.³ Finally, a dynamic equilibrium is simply a sequence of static equilibria linked through bequest decisions.

We will now analyze the equilibrium of the economy under two scenarios:

1. Effort choices of managers are publicly observed (the case of perfect information).
2. Effort choices of managers are not observed by any other agent in the economy (the case of imperfect information). In this case firms will have to obey the incentive compatibility constraints of their managers.

2.3 Equilibrium With Perfect Information

Since output in this economy will be non-random and markets are complete, risk-neutral firms will pay non-random wages, managerial salaries and interest rates. Therefore, $w_j(\omega_t; \mathbf{M}_t, K_t) = w_j(\mathbf{M}_t, K_t)$, and similarly for z_{jt} and r_t . Moreover, all agents in the same occupation in island j will receive the same payment, and all firms in island j will adopt the same technology and hire the same amount of capital and labor. Further because agents are free to become managers or workers, all agents within an island must be indifferent between these two occupations. Therefore, $z_j(\mathbf{M}_t, K_t) = w_j(\mathbf{M}_t, K_t) + \frac{1}{2\beta}e_{jt}^2$ where e_{jt} is the effort exerted by each manager in island j , and $z_j(\mathbf{M}_t, K_t)$ is the salary of the manager conditional on exerting the agreed level of effort, e_{jt} .

Before characterizing the equilibrium occupational choice in each island, we also assume that

$$K_t \geq N^{-\frac{\alpha}{1-\alpha}}, \quad (6)$$

³It is however possible in theory that two firms in island j offer different state-contingent functions, say $w_{jt}(\omega)$ and $w'_{jt}(\omega)$ but workers are indifferent between these two functions and firms make the same expected profits. Leaving this case out is without loss of generality, since it does not happen in equilibrium, and enables us to keep the definition of equilibrium relatively simple.

holds at time t . This condition ensures that there is enough capital in the economy at time t so that if this capital is allocated equally across all the islands, at least one firm in each island can be run with *productive efficiency*, namely adopting a technology such that $l_{jt}^\alpha k_{jt}^{1-\alpha} = 1, \forall j$. When (6) is satisfied, all firms are run with productive efficiency and fully utilize all of their managerial input. We can now state the following Lemma.⁴

Lemma 1 *Suppose (6) holds at time t and there is perfect information. Then, in equilibrium:*

1. $e_{jt} = e^{fb} = \beta$ in all islands;
2. the equilibrium is symmetric in the sense that capital is equally allocated across islands (for all $j \in [0, 1]$, $K_{jt} = K_t$). The number of firms (managers) which are active in each island is

$$M_t = M(K_t) \equiv N^\alpha K_t^{1-\alpha}. \quad (7)$$

All firms hire the same amount of labor and capital: $l(K_t) = \left(\frac{N}{K_t}\right)^{1-\alpha}$ and $k(K_t) = \left(\frac{K_t}{N}\right)^\alpha$ so that $l_t^\alpha k_t^{1-\alpha} = 1$;

3. the economy-wide interest rate, and the wage rate and managerial compensation in every island $j \in [0, 1]$ are given by:

$$r(\omega_t; \mathbf{M}_t, K_t) = r(K_t) = (1 - \alpha) \left(\frac{N}{K_t}\right)^\alpha \left(\mu + \frac{\beta}{2}\right); \quad (8)$$

$$w_j(\omega_t; \mathbf{M}_t, K_t) = w(K_t) = \alpha \left(\frac{K_t}{N}\right)^{1-\alpha} \left(\mu + \frac{\beta}{2}\right); \quad (9)$$

$$z_j(\omega_t; \mathbf{M}_t, K_t) = z(K_t) = \alpha \left(\frac{K_t}{N}\right)^{1-\alpha} \left(\mu + \frac{\beta}{2}\right) + \frac{\beta}{2}. \quad (10)$$

Lemma 1 establishes a number of results. First, since there is no informational imperfections, managers will agree to exert the first-best level of effort, $e^{fb} = \beta$, which equates marginal cost to the marginal benefit of higher return (part 1). Second, decreasing returns to capital ensures that the equilibrium is symmetric in the sense that all islands receive the same

⁴Throughout the paper we ignore integer problems and use differential calculus with M_j .

amount of capital (part 2). Note that from (7), M_t is increasing in K_t , thus as the amount of capital in the economy expands, the number of managers (firms) increases. As a result, development is associated with capital deepening and a growing proportion of the agents who choose the managerial occupation. Finally, although output in any particular island is random, thanks to the large number of islands, total output in this economy is non-random: $Y_t = \int_0^1 (\mu + e_{jt}) \cdot \min[1, l_{jt}^\alpha k_{jt}^{1-\alpha}] M_{jt} dj = (\mu + \beta) N^\alpha K_t^{1-\alpha}$. Together with the lack of informational imperfection, this ensures that risk-neutral firms can offer full insurance to the factors of production that they hire (part 3). In particular, (8) and (9) state that in equilibrium the expected revenue of each firm net of the additional cost of managerial compensation, $(\mu + \frac{\beta}{2})$, is distributed to capital and labor with shares $(1 - \alpha)$ and α . (10) ensures that managers get exactly the same expected return as workers.

Given Lemma 1, equilibrium dynamics are straightforward. Since a fraction s of all earnings are saved, the equilibrium capital stock follows the recursion:

$$K_{t+1} = sY_t = s(\mu + \beta) M(K_t) \quad (11)$$

$$= s(\mu + \beta) N^\alpha K_t^{1-\alpha}. \quad (12)$$

We also assume:

$$s(\mu + \beta) > N^{-\frac{\alpha}{1-\alpha}} \quad (13)$$

This condition – which is always satisfied when N is sufficiently large – ensures that the steady state level of capital is large enough so that more than one firm per island can be opened, and guarantees that the capital stock of the economy does not fall below a certain lower bound. Then :

Proposition 1 *Assume that (6) holds at $t = 0$, that (13) is satisfied, and that there is perfect information. Then, there is a unique equilibrium sequence of allocations where within every period, the number of firms, wages and managerial salaries in each island, and the interest rate in the economy are given by Lemma 1, and the aggregate capital stock, K_t , follows (11). K_t uniformly converges to the unique steady-state capital stock, $K^{ss} = [s(\mu + \beta)]^{1/\alpha} N$.*

The equilibrium dynamics of the economy under perfect information are neoclassical: there is accumulation until a steady state is reached and the rate

of return on capital decreases monotonically in the process. Furthermore, given the absence of informational asymmetries, neither the variability of rewards nor the power of incentives change over time. As a result, the behavior of managers and the organization of production are independent of the stage of development.

3 The Economy With Imperfect Information

We will now assume that the effort choice of a manager is his private information. This introduces standard moral hazard considerations and implies that the manager should be rewarded conditional upon his performance, and thus will have to receive a random return. We also assume that while the ex-post performance of each individual firm can be costlessly observed, neither the island-specific (a_j) nor firm-specific (ϵ_{ij}) productivity shocks are publicly observed.

We will first characterize the equilibrium wages, the rate of return to capital, and the form of equilibrium managerial contracts conditional upon the allocation of the capital stock to different islands. We will next show that under certain conditions, only a unique symmetric equilibrium will exist, and we will fully characterize the dynamic equilibrium in this case. We will end this section with some comparative static results and a discussion of constrained efficiency.

3.1 Static Equilibrium

Let us first define $\zeta_j(\omega_t; \mathbf{M}_t, K_t) \equiv z_j(\omega_t; \mathbf{M}_t, K_t) - w_j(\omega_t; \mathbf{M}_t, K_t)$ as the compensation that a manager in island j receives additional to the wage component of his earnings. ζ_j will compensate him for the effort cost and for the risks he takes (in the case of perfect information, we had $\zeta_j = \frac{\beta}{2}$).

Lemma 2 *Suppose (6) holds at time t . Then, in equilibrium:*

1. *for all $j \in [0, 1]$, the number of firms (managers) which are active in island j is given by $M_j(K_{jt}) = M_{jt}$ such that:*

$$M_{jt} = N^\alpha K_{jt}^{1-\alpha} \text{ and } \int_0^1 K_{jt} dj = K_t. \quad (14)$$

All firms in island j hire the same amount of labor and capital: $l_{jt} = l(K_{jt}) = \left(\frac{N}{K_{jt}}\right)^{1-\alpha}$ and $k_{jt} = k(K_{jt}) = \left(\frac{K_{jt}}{N}\right)^\alpha$ so that $l_{jt}^\alpha k_{jt}^{1-\alpha} = 1$;

2. the economy wide interest rate and wages in every island $j \in [0, 1]$ are given by:

$$r(\omega_t; \mathbf{M}_t, K_t) = r(\mathbf{M}_t, K_t) = \frac{1-\alpha}{k(K_{jt})} E_t[y_{jt} - \zeta_j(\omega_t; \mathbf{M}_t, K_t)], \quad (15)$$

$$w_j(\omega_t; \mathbf{M}_t, K_t) = w_j(\mathbf{M}_t, K_t) = \frac{\alpha}{l(K_{jt})} E_t[y_{jt} - \zeta_j(\omega_t; \mathbf{M}_t, K_t)] \quad (16)$$

where E_t is the expectations operator conditional on the public information set of time t .

Part 1 of Lemma 2 is identical to part 2 of Lemma 1, except that under imperfect information the equilibrium does not necessarily have capital equally invested in all islands. Part 2 of Lemma 2, the analogue of part 3 of Lemma 1, states that there is no issue of risk-taking by labor and capital (hence r and w_j do not depend on the state of nature ω_t). The large number of islands ensures that there is no aggregate risk, thus risk-neutral firms can once more offer full insurance to the factors for which there is no incentive problem. Therefore, as in the case of perfect information, the expected revenue of firms net of managerial premium will be distributed between capital and labor, again with shares α and $1 - \alpha$. In contrast to returns to capital and labor, managerial compensation will be random because individual managers will have to bear some risk to provide them with the right incentives. The rest of this section will characterize the contract which determines the managerial compensation.

We start the analysis with two observations. First, since the economy has a linear structure, normally distributed random variables, and CARA utility, we can appeal to the results of H-M (1987) who prove that with this structure and continuous adjustment of effort levels, the optimal contract is linear. Moreover, thanks to CARA preferences and normally distributed returns, we can simply work with the *certainty equivalent* of the income process faced by the agents (see also H-M, 1991, for an application). Even though we have so far thought of effort as chosen once and for all, our structure would carry over unchanged if we considered each period to be a segment of continuous time, and managers continuously adjusted their effort after observing previous

performance. Using the result of Holmstrom and Milgrom and this slightly modified interpretation of our timing structure, we will restrict attention to linear contracts. Second, we know from standard agency theory that any variable which contains information about the effort level will be useful in giving incentives to the agent (Holmstrom, 1979). In our economy, average output in the island of the manager in question will be a useful variable to condition contracts upon. Average output of island j is correlated with a_{jt} , and conditioning on a_{jt} is beneficial because the variability generated by this shock reduces the power of incentive contracts. To see the intuition, imagine that firm i in island j performed very poorly. If all other firms in the island did well, this suggests that the island must have received a favorable shock and the bad performance of the manager is likely to have been due to low effort. In contrast, if all other firms affected by the same shock also performed badly, it is likely that the poor performance was due to an adverse island specific shock, not to low effort.

Let us now drop time subscripts whenever this will cause no confusion. The optimal compensation contract for the manager of firm i in island j will then take the form:⁵ $z_{ij} = \hat{\phi}_{0ij} + \phi_{1ij}(y_{ij} - \mu) + \phi_{2ij}(y_{ij} - y_{j(-i)}^a)$ where $y_{j(-i)}^a$ is the average productivity of all j -th island firms except firm i , i.e. $y_{j(-i)}^a = \frac{\sum_{h \neq i} y_{hj}}{M_j - 1}$. Rewriting in terms of the compensation of the manager additional to the wage, ζ_{ij} (recall $z_{ij} = w_j + \zeta_{ij}$):

$$\zeta_{ij} = \phi_{0ij} + \phi_{1ij}(y_{ij} - \mu) + \phi_{2ij}(y_{ij} - y_{j(-i)}^a) \quad (17)$$

where $\phi_{0ij} \equiv \hat{\phi}_{0ij} - w_j$. Note that the compensation of the manager is made conditional upon the performance of the firm (the term $y_{ij} - \mu$), and the relative performance compared to the average output of all other firms in the same island (the term $y_{ij} - y_{j(-i)}^a$). Expressed differently, (17) will be a type of *relative performance evaluation* contract which sets the average performance of other agents as the *standard* relative to which the manager is judged. Quite importantly, as the number of firms in island j , M_j , increases, $y_{j(-i)}^a$ will be more strongly correlated with a_{jt} , and the standard that the society can set

⁵See Theorem 7 of H-M (1987). Note that with this contract, the manager will sometimes have to receive a negative payment. However, for μ large enough this will happen very seldom, and moreover, since there is accumulated wealth, negative payments are not problematic. In what follows, we ignore the constraint that wealth should be non-negative.

will become more accurate.

Next, it can be shown that productive inefficiency is never compatible with equilibrium as long as there are at least two active firms in each island, and thus we will limit our attention to equilibria with productive efficiency (as defined in section 2.3). Then, the problem of firm i on island j is equivalent to solving the following program:

$$\max_{\phi_{0ij}, \phi_{1ii}, \phi_{2ij}} E(y_{ij} - \zeta_{ij} - w_j l_j - r k_j \mid e_{ij} = e_{ij}^*) \quad (18)$$

subject to:

$$e_{ij}^* = \arg \max_{e_{ij}} E(\zeta_{ij}) - \frac{1}{2\beta} e_{ij}^2 - \frac{\rho}{2} \text{Var}(\zeta_{ij}) \quad (19)$$

$$E(\zeta_{ij} \mid e_{ij}^*) - \frac{\rho}{2} \text{Var}(\zeta_{ij}) = \frac{1}{2\beta} e_{ij}^{*2} \quad (20)$$

where y_{ij} is given by (3) and ζ_{ij} is as in (17). (18) is clearly the expected profit of the firm. The first condition, (19), is the incentive compatibility constraint. It requires that the effort choice of the agent should be the one that maximizes his payoff given the managerial contract. Note that we have made use of the fact that the manager will simply maximize the certainty equivalent of his income minus the cost of effort. (20) is the participation constraint which requires that the certainty equivalent of the additional return (over and above the wage w_j) that the manager receives should exactly compensate him for the cost of effort. The participation constraint (20) is *necessary* and *sufficient* to characterize the equilibrium occupational choices. To see why, recall that there are no wealth effects, thus if one agent prefers to be a manager rather than a worker, all agents in the same island would also prefer to become managers irrespective of their wealth level. Therefore, in equilibrium all agents must be indifferent between the two occupations and (20) ensures this. Also, note that in the objective function (18), we have imposed that each firm takes the price of capital, r , and the price of labor in island j , w_j , as given. This is to be understood as each firm taking the capital stock of the economy, K_t , and the first-stage announcements of all other firms, \mathbf{M}_t , as given and anticipating the equilibrium price of capital and the wage rate in island j in the second stage of the entry game (see section 2.2).

CARA preferences together with linear contracts simplify the problem, allowing us to proceed in two steps. In particular, because utility is transferable,

we can first maximize the sum of the firm's and the manager's utility with respect to ϕ_{1ij} and ϕ_{2ij} subject to the incentive compatibility of the manager, (19). Ignoring terms which do not affect the solution, this maximization problem can be written as:

$$\max_{\phi_{1ij}, \phi_{2ij}} E(y_{ij} | e_{ij}^*) - \frac{1}{2\beta} e_{ij}^{*2} - \frac{\rho}{2} \text{Var}(\zeta_{ij}) \quad (21)$$

subject to (19). Next, ϕ_{0ij} can be determined by solving the participation constraint, (20). The following Lemma establishes three important intermediate results.

Lemma 3 *Under imperfect information:*

1. *effort choice of manager i in island j is given as:*

$$e_{ij}^* = \beta (\phi_{1ij} + \phi_{2ij}). \quad (22)$$

2. *the average productivity of firm i in island j is: $E(y_{ij} | e_{ij}^*) = (\mu + e_{ij}^*)$.*
3. *the variance of managerial compensation for manager i is given by:*

$$\text{Var}(z_{ij}) = \text{Var}(\zeta_{ij}) = \left[(\phi_{1ij} + \phi_{2ij})^2 \sigma^2 + \phi_{1ij}^2 \nu^2 + \phi_{2ij}^2 \frac{\sigma^2}{M_j - 1} \right]. \quad (23)$$

Lemma 3 enables us to fully characterize the set of equilibrium contracts.

Proposition 2 *Suppose (6) holds. Then in equilibrium all managers in island j have contracts: $\zeta_{ij}^* = \phi_{0j}^* + \phi_{1j}^* (y_{ij} - \mu) + \phi_{2j}^* (y_{ij} - y_{j(-i)}^a)$, (with $z_{ij}^* = w_j + \zeta_{ij}^*$) where:*

$$\phi_{1j}^* = \phi_1^*(K_j) = \frac{\sigma^2}{(M_j \nu^2 + \sigma^2) \frac{\rho \sigma^2}{\beta} + \sigma^2 + (M_j - 1) \nu^2}, \quad (24)$$

$$\phi_{2j}^* = \phi_2^*(K_j) = \frac{(M_j - 1) \nu^2}{(M_j \nu^2 + \sigma^2) \frac{\rho \sigma^2}{\beta} + \sigma^2 + (M_j - 1) \nu^2}, \quad (25)$$

$$\phi_{0j}^* = \phi_0^*(K_j) = \frac{\beta}{2} (\phi_{1j}^* + \phi_{2j}^*)^2 + \frac{\rho}{2} \text{Var}(\zeta_j^*). \quad (26)$$

and $\text{Var}(\zeta_j^*)$ is given by (23) with $\phi_{1ij} = \phi_{1j}^*$ and $\phi_{2ij} = \phi_{2j}^*$.

This proposition establishes that all firms in island j will choose exactly the same managerial contract and this is uniquely determined for given M_j (which itself is a function of K_j from (14))⁶ The dependence of both ϕ_{1j}^* and ϕ_{2j}^* on M_j implies that as the number of firms in island j increases incentive contracts change. Also, since $e_j^* = \beta (\phi_{1j}^* + \phi_{2j}^*)$ (Lemma 3), the organization of production – here captured solely by the level of managerial effort – depends on the number of firms in the island, M_j . The comparative statics in section 3.4 will establish that $\frac{de_j^*}{dM_j} > 0$. Intuitively, as commented above, when there are more firms, the society can set *more accurate standards*; this enables a change in managerial contracts and an increase in effort and productivity. Combining the results of Proposition 2 with those of Lemma 2, we can easily obtain equilibrium factor returns conditional on \mathbf{M}_t , i.e. on the allocation of capital across islands. We next discuss this allocation.

3.2 Symmetric vs. Asymmetric Equilibria

In the perfect information case, we established that there could only be a *symmetric* equilibrium whereby all islands receive the same amount of capital, $K_{jt} = K_t$ (and $M_{jt} = M_t$). With imperfect information, relative performance evaluation and the endogeneity of information introduce an island-specific *externality*, and as a result, the equilibrium may involve an asymmetric distribution of the total capital stock across the islands. More specifically, information and incentives improve with the scale of production within an island and this counteracts decreasing returns to capital in the island. In an asymmetric equilibrium, islands which receive a higher (lower) than average amount of investments have higher (lower) capital to labor ratios, and this depresses (increases) the rate of return to capital. But at the same time, the larger (smaller) number of firms improves (reduces) information and productivity, and this increases (reduces) the rate of return to capital.

⁶As we will see in more detail later, this feature implies that when the number of firms is larger in island j information problems are less severe. This may suggest that it could sometimes be profitable to increase the number of firms by sacrificing productive efficiency. However, the form of our production function (3) precludes this possibility. What matters is not the number of other firms producing in the same island, but total production. M_j appears in our expressions because when all firms are run with productive efficiency, total output is proportional to M_j . This also explains the particular form of the production function chosen.

Lemma 4 below establishes the uniqueness of a symmetric equilibrium under simple restrictions.

Lemma 4 *$\exists \bar{\mu}$ such that, for all K_t satisfying (6), $\forall \mu \geq \bar{\mu}$, there exists a unique equilibrium which is symmetric.*

We can see the intuition for this lemma as follows. μ is the amount of output each firm produces irrespective of the effort level of the manager. The allocation which maximizes the number of firms in the economy has capital allocated equally across islands. Thus, μ is the opportunity cost of allocating capital asymmetrically (i.e. of reducing the number of firms in the economy). As a consequence, when μ is large, the opportunity cost of an asymmetric distribution of capital is prohibitively high, and there only exists a unique symmetric equilibrium.

3.3 Dynamics of Symmetric Equilibria

Since asymmetric equilibria are complicated and of only indirect interest for the focus of this paper, in this section we will fully characterize only the dynamics of accumulation and development with symmetric equilibria. We will discuss asymmetric equilibria in section 4.6.

In the case of a symmetric equilibrium, managers in all islands will receive exactly the same contract, $\phi_{0jt}^* = \phi_0^*(K_t)$, $\phi_{1jt}^* = \phi_1^*(K_t)$, $\phi_{2jt}^* = \phi_2^*(K_t)$, where $\phi_0^*(K_t)$, $\phi_1^*(K_t)$ and $\phi_2^*(K_t)$ are given by Proposition 2 and simply depend on K_t since in a symmetric equilibrium $M_{jt} = N^\alpha K_t^{1-\alpha}$. Furthermore, given Proposition 2, all firms in the economy adopt the same technology, and workers in different islands receive the same wage.

Since a fraction s of all income is saved, the law of motion of capital with symmetric equilibria can be written as:

$$K_{t+1} = s [\mu + \beta (\phi_1^*(K_t) + \phi_2^*(K_t))] N^\alpha K_t^{1-\alpha}. \quad (27)$$

We can summarize our findings in (proof in the text):

Proposition 3 *Assume that the conditions of Lemma 4 and condition (6) are satisfied. Then, given K_t , there exists a unique symmetric equilibrium allocation at time t . In this equilibrium $M_{jt} = N^\alpha K_t^{1-\alpha}$ for all $j \in [0, 1]$, and*

all managers sign contract (17) with $\phi_0^*(K_t)$, $\phi_1^*(K_t)$ and $\phi_2^*(K_t)$ as given by Proposition 2, and choose effort level as in (22). Factor prices are:

$$\begin{aligned} r(K_t) &= (1 - \alpha) (\mu + \beta (\phi_1^*(K_t) + \phi_2^*(K_t)) - \phi_0^*(K_t)) \left(\frac{N}{K_t} \right)^\alpha \\ w(K_t) &= \alpha (\mu + \beta (\phi_1^*(K_t) + \phi_2^*(K_t)) - \phi_0^*(K_t)) \left(\frac{K_t}{N} \right)^{1-\alpha} \end{aligned} \quad (28)$$

The evolution of the physical capital stock is given by (27).

This proposition establishes that in the dynamic equilibrium of our economy, capital accumulation is accompanied by an increase in the number of firms and more repetition. The information that is accumulated as a result of this process enables the society to set better standards, and improves the effort level of managers and total factor productivity. Thus, the interaction of endogenous information and incentives creates a form of “scale externality”. Because total factor productivity may be increasing in the capital stock over a certain range, dynamics are no longer purely neoclassical, and multiple interior stable steady states cannot be ruled out in general (though it can be established that for μ sufficiently large, the steady state is unique).

Remark We have so far talked of islands. For some of our applications an economy consisting of different sectors may be more appropriate. To start with, our results would apply exactly to an economy where there is a continuum of sectors, each agent has a strong comparative advantage for one sector and the output of different sectors are perfect substitutes. A more realistic formulation would involve different sectors producing *imperfect* substitutes. In this case, aggregate consumption could be defined as a composite of different sectors’ output, e.g. $c_t = \exp \left[\int_0^1 c_{jt} dj \right]$, and agents could be homogeneous and decide which sector to work in. This setup – which was analyzed in a previous version of this paper – gives similar results, but the analysis is more involved due to “Jensen’s inequality” terms in aggregation.

3.4 Some Comparative Statics

Equations (24) and (25) lead to a number of interesting comparative static results. Straightforward differentiation establishes:

$$\begin{aligned}
\frac{d\phi_1^*}{d\sigma^2} &< 0, & \frac{d\phi_2^*}{d\sigma^2} &< 0, & \frac{de^*}{d\sigma^2} &< 0; \\
\frac{d\phi_1^*}{d\nu^2} &< 0, & \frac{d\phi_2^*}{d\nu^2} &> 0, & \frac{de^*}{d\nu^2} &> 0; \\
\frac{d\phi_1^*}{dM} &< 0, & \frac{d\phi_2^*}{dM} &> 0, & \frac{de^*}{dM} &> 0.
\end{aligned} \tag{29}$$

First, when σ^2 increases, both ϕ_1^* and ϕ_2^* decrease, and effort and productivity fall. To understand this result, recall that in this economy it is precisely idiosyncratic variability which makes it costly to induce effort. If $\sigma^2 = 0$, managerial contracts would specify $\phi_1^* = 0$ and $\phi_2^* = 1$, and provided that there are at least two firms in the island, managers would bear no risk, and the first-best would be implemented. Idiosyncratic variability introduces noise to the signal coming from the individual performance and makes managerial compensation random. Since managers are risk-averse, this lack of full insurance is costly, and managerial contracts trade off insurance and incentives. As σ^2 increases, lack of insurance becomes more costly, and there is lower effort in equilibrium. Second, when ν^2 increases, ϕ_1^* falls and ϕ_2^* increases because with more island specific variability, relative standards become more informative. Overall, the change in ϕ_2^* dominates and the net effect is that the level of effort and productivity increase with the volatility of island-specific shocks. Third, as the number of firms grows (as a result of capital accumulation), the standards improve, and again there is more of relative and less of absolute performance evaluation (higher ϕ_2^* and lower ϕ_1^*). Once more the effect through ϕ_2^* dominates, and we have $\frac{de^*}{dM} > 0$, thus accumulation of information (and capital) leads to higher managerial effort and productivity. This contrasts with the economy with perfect information where there was no interaction between incentives and development.

3.5 Agency Costs and Development

Agency costs are the costs incurred by the society due to the imperfect information in principal-agent relations. In our model these have two components: (i) managers exert less effort in the economy with asymmetric information than in the first best; and (ii) they require a risk-premium to be compensated for the variability in their income. We capture both components with our

concept of $TAC(K)$, *total agency costs* that the society incurs *per firm*:

$$TAC(K) \equiv \left[(\beta - e^*) - \frac{1}{2\beta} (\beta^2 - (e^*)^2) \right] + \frac{\rho}{2} Var(\zeta^*) \quad (30)$$

where e^* and $Var(\zeta^*)$ depend on $M(K)$, $\phi_1^*(K)$ and $\phi_2^*(K)$ via equations (14), (22) and (23), and this makes TAC depend on K . The first term of (30) is the effort component and the second is the loss of utility in certainty equivalent terms due to the risk borne by the managers.

Another useful concept is $SAC(\bar{e}, K)$ (*shadow agency cost*) which is given by the certainty equivalent of income that is foregone in order for managers to be induced to exert the effort level \bar{e} (as different from the optimal level of effort, e^*). Formally:

$$SAC(\bar{e}, K) \equiv \min_{\{\phi_1, \phi_2\}} \left[\frac{1}{2\beta} \bar{e}^2 + Var(\zeta) \right] \quad s.t. \quad \phi_1 + \phi_2 = \frac{\bar{e}}{\beta} \quad (31)$$

where $Var(\zeta)$ is given by (23).

Proposition 4 *Both $TAC(K)$ and $SAC(\bar{e}, K)$ are weakly decreasing functions of the capital stock of the economy.*

This proposition establishes that more information is always useful in our setting as it enables better incentives and risk sharing. Therefore, the cost of more effort at the margin and the total loss of utility due to incentive problems are decreasing in the amount of information, and thus in the total capital stock of the economy.

3.6 Constrained Efficiency

We have established that as the economy develops, it achieves higher levels of effort and productivity. Can a social planner subject to the same technological and informational constraints ensure a more efficient outcome? To answer this question, we analyze the static problem of a planner maximizing the sum of the utility of all agents in the economy without any distributional concern.⁷ We start with three simple observations. First, as in the decentralized economy, as long as there is enough capital to open at least two firms with productive efficiency in every island, the planner would never choose productive inefficiency.

⁷We ignore saving decisions which will depend on the “discount rate” of the planner.

Second, given our assumptions, the planner will also choose linear contracts. Finally, the planner will offer the same contracts to all managers in the same island. Then, the planning problem can be written as:

$$\max_{\{\phi_{1j}, \phi_{2j}, M_j, K_j, e_j\}} \int_0^1 M_j (\mu + e_j) dj - \frac{\rho}{2} \int_0^1 M_j Var(\zeta_j) dj - \frac{1}{2\beta} \int_0^1 M_j e_j^2 dj \quad (32)$$

subject to the incentive compatibility of the managers, (19), and the resource constraint (14), with $Var(\zeta_j)$ given by (23).

Proposition 5 *Conditional on \mathbf{M}_t , the planner chooses $\phi_{1j}^s = \phi_{1j}^*$, $\phi_{2j}^s = \phi_{2j}^*$ and induces $e_j^s = e_j^*$ for all $j \in [0, 1]$ as given by equations (22), (24) and (25).*

The proposition states that conditional on the allocation of capital across islands, the planner would choose the same allocation as the decentralized economy, or in other words, she would choose exactly the same contracts and induce the same level of effort. Although there are many externalities at work, the contracts that result from decentralized competition are efficient. To understand the intuition of this result, first note that the effort level of a manager does *not* create an externality on other managers in the same island. Given the additive structure of (3), as long as he exerts the effort level he is expected to (a requirement in any pure strategy equilibrium), the signal extraction problem faced by all other firms is unaffected. This can also be seen in equations (24) and (25) where the power of incentives given to the manager only depends on the total amount of production (or equivalently, on the number of firms) in the same island.

Despite the fact that contract choices are efficient, the allocation of capital across islands chosen by the planner does not necessarily coincide with the equilibrium. More specifically, it is straightforward to show, along the lines of Lemma 4, that if μ is sufficiently large (say greater than $\bar{\mu}^S$), the planner will choose a symmetric equilibrium. Therefore, when μ is larger than both $\bar{\mu}$ and $\bar{\mu}^S$, the unique symmetric equilibrium that we characterized above is also the constrained efficient allocation. However, it is not possible to establish unambiguously how $\bar{\mu}^S$ compares to $\bar{\mu}$. Intuitively, when a firm decides to locate in island j rather than j' , it ignores two externalities it is creating: workers in island j will be better off and those in island j' will be worse off

because labor demand and wages will increase in island j and decrease in island j' . These two effects do not always cancel out in the equilibrium, hence the distribution of capital across islands is not necessarily efficient. Moreover, it is useful to note at this stage that many of the applications we will consider in the next section features technologies or sectors with different degrees of agency costs, and in these situations, the decentralized equilibrium is more likely to be inefficient.

4 Applications

4.1 Risk Sharing

Since workers and capital owners bear no risk, the extent of risk sharing in our economy is captured by the variance of managerial returns $Var(\zeta(M(K), \phi_1^*(K), \phi_2^*(K)))$. This is a component of total agency cost (see (30)) analyzed in the previous section. However, although $TAC(K)$ decreases with accumulation, this is not necessarily true for the variance. That is, the degree of risk sharing may be non-monotonic or even decreasing with development.

Proposition 6 *Let $V(K) \equiv Var(\zeta(M(K), \phi_1^*(K), \phi_2^*(K)))$. Then:*

1. *if $\rho(\sigma^2 + \nu^2) < \beta$, then $V'(K) < 0$ for all K ;*
2. *if $\rho\sigma^2 < \beta < \rho(\sigma^2 + \nu^2)$, then $\exists \bar{K}$ s.t. if $K < \bar{K}$ then $V'(K) > 0$, and if $K > \bar{K}$ then $V'(K) < 0$;*
3. *if $\rho\sigma^2 > \beta$, then $V'(K) < 0$ for all K .*

Intuitively, as M (i.e. K) increases, shadow agency cost, $SAC(\bar{e}, K)$, falls but in the mean time, the equilibrium level of effort e^* also increases, and this requires managers to bear more risk. This interaction between two opposing forces determines how the variability of managerial returns will change over the development process. Proposition 6 shows that the link between risk sharing and growth depends on the degree of risk-aversion and on the amount of noise that contractual arrangements are subject to. For instance, if agents have a low degree of risk-aversion (small ρ) or the variance of the shocks is small, then, as more information becomes available, risk sharing will improve. The opposite

occurs when the degree of risk aversion (or the idiosyncratic variability, σ^2) is high. In this case, because incentives are very low powered, the variability of managerial returns is very low in poor economies, and increases over time. In intermediate cases, the variance is non-monotonic, and risk sharing increases first, and decreases afterwards. Even though only managers bear risks in our economy, the opposing forces impacting on risk sharing will apply more generally to all agents bearing risks due informational problems.

The possibility that risk sharing is decreasing with accumulation (case 3), or inverse U-shaped (case 2), provides an interesting interpretation to some recent empirical evidence. It is often argued that less developed economies suffer from serious agency problems (see for instance, North, 1990). However, Townsend (1994, 1995a,b) and other recent studies (as reviewed by Morduch, 1995) have shown that in Asian villages there is relatively low variance of consumption, thus quite good risk sharing arrangements. Moreover, Townsend (1995b) finds that risk sharing appears to be *lower* in *richer* villages. It is tempting to interpret these recent findings as evidence that less developed economies do not in fact suffer serious incentive problems, and that growth and modernization may be the factors destroying the “efficient” organization of small communities and villages. Our model provides an alternative interpretation to these findings whereby in less developed economies, the organization of production is highly inefficient because *shadow agency costs* are prohibitively high. As the scale of economic activity grows, more information becomes available, and more sophisticated contractual arrangements can be devised. Because these contracts induce higher effort, the extent of observed risk sharing may decline. Possibly at even later stages of development, the variability of managerial and entrepreneurial returns may again start falling as relative performance evaluation becomes more powerful.

A related feature worth noting is that our model also has implications about the distribution of (labor) income. Income distribution is determined by the choice of effort by managers and the variability of managerial returns. In cases 2 and 3 of Proposition 6, growth is “unequalizing” in less developed economies because it leads to an increase in managerial effort and higher variability of managerial incomes, and therefore to a greater difference between the average income of managers and workers. In case 2, however, as capital accumulates further, the variability of managerial returns will decrease, and this will reduce

both the observed dispersion among entrepreneurs and the risk-premium that managers are paid over workers. If the decline in risk-premia dominates the increased compensation for higher effort, growth will bring about a more equal distribution of income in advanced economies. Therefore, our model with intermediate levels of risk aversion is consistent with the Kuznets curve.

4.2 Sectorial Transformations

Sectors typically differ in terms of the structure of uncertainty they face as well as their technology. This will imply that agency problems will be more serious in certain sectors than others. When the scale of production is limited, there will be very little information to be used in agency relations, and sectors where agency matters more will have relatively lower productivity. As a result, capital accumulation will be accompanied by sectorial transformation towards activities where agency problems are more important. Even though this insight may have a number of potential applications, in this section we will consider the implications of our model for how the share of agriculture in total production evolves over the process of development.

As discussed above, the variance of idiosyncratic shocks relative to the variance of common shocks are crucial for the extent of agency costs. Now suppose, not implausibly, that agriculture is subject to large common shocks due to weather, while the variation due to idiosyncratic uncertainty (or managerial talent) are more important in industry (manufacturing). This assumption is in accord with the empirical evidence provided by Pollard (1965) who documents the very high failure rates of managerial firms during the seventeenth and eighteenth century Britain, and attributes most of this to managerial mistakes or negligence, and with Townsend (1995b) who reports that individuals involved in entrepreneurial activities suffer more volatile consumption than farmers.

Let us now develop a very simple version of this sectorial transformation story by making three strong assumptions: **(i)** There are two goods. We think of the first as an agricultural and the second as a manufacturing (industrial) product. Half of the islands can only produce agricultural goods and the other half can only produce industrial goods. On each island $j \in [0, 1]$ there are N agents which can only work in agriculture, while on each island $j \in [1, 2]$ there are N agents which can only produce industrial goods. **(ii)** Agricultural and

industrial products are perfect substitutes. **(iii)** The variance of idiosyncratic shocks in agriculture is $\sigma_A^2 = 0$ and the variance for the idiosyncratic shocks in manufacturing is $\sigma_I^2 > 0$. Thus, agency problems are absent in “agriculture”.⁸ However, note that since $\nu_A^2 > 0$, agricultural output may be subject to more variability than manufacturing.

The technology is essentially the same as in the one-sector economy. All firms have a quasi-Leontieff technology like (3). Workers cannot move across islands but can invest their wealth in a balanced portfolio of all the firms in the economy, thus will bear no risk. Since in equilibrium all firms are run under *productive efficiency*, we have:

$$\begin{aligned} y_{ijt}^A &= Z \left(\mu + e_{ijt}^A + a_{jt}^A \right) \quad \forall j \in [0, 1], \\ y_{ijt}^I &= \mu + e_{ijt}^I + \epsilon_{ijt}^I + a_{jt}^I \quad \forall j \in [1, 2], \end{aligned} \quad (33)$$

where Z measures the productivity of agriculture relative to industry. Furthermore:

$$a_j^A \sim N(0, \nu_A^2), \quad a_j^I \sim N(0, \nu_I^2), \quad \epsilon_{i,j}^I \sim N(0, \sigma_I^2). \quad (34)$$

We will also assume, in analogy with the result of Lemma 4, that parameters are such that within each sector there is only a symmetric equilibrium, i.e. $M_j^A = M^A$ and $M_j^I = M^I$ for all $j \in [0, 1]$, though in general $M^A \neq M^I$. Moreover, managerial contracts in agriculture z^A (or ζ^A) will differ from managerial contracts in industry z^I (or ζ^I) because of the differences in the structure of uncertainty. Consequently, managerial effort in agriculture, e^A , will differ from managerial effort in industry, e^I . In particular, since the return to agricultural firms within each island are perfectly correlated, the first-best effort level can be implemented in agriculture, that is $e^A = Z\beta$ and $Var(z^A) = 0$. Instead, industrial contracts will induce the effort level $e^I = \beta(\phi_1^* + \phi_2^*)$, with ϕ_1^* and ϕ_2^* given by (24) and (25). Note that the contract in the industrial sector will be conditional on the information – thus, the number of firms – in the industrial islands.

⁸All three assumptions can be relaxed. For instance, islands can be allowed to choose whether to specialize in agriculture or industry. Instead of perfect, the two goods may be imperfect substitutes with elasticity of substitution greater than one, and this would also enable us to match the relative price movements over the development process, but again is not crucial for our argument. Also, $\frac{\sigma_I^2}{\nu_I^2} > \frac{\sigma_A^2}{\nu_A^2}$ rather than $\sigma_A^2 = 0$ would be sufficient in general.

Let us now write the rate of return on capital in the two sectors, r^A and r^I , when firms make zero profits and labor is paid its marginal product. As in the previous sections, these are given as:

$$r^A = \frac{(1 - \alpha)Z \left(\mu + \frac{Z\beta}{2} \right)}{k^A} \quad (35)$$

$$r^I = \frac{(1 - \alpha) \left(\mu + e^I - \frac{(e^I)^2}{2\beta} - \frac{\rho}{2} \text{Var}(\zeta^I) \right)}{k^I} = \frac{(1 - \alpha) \left(\mu + \frac{\beta}{2} - TAC^I(M^I) \right)}{k^I} \quad (36)$$

where the second equality of (36) follows from defining TAC^I as the total agency costs in industry analogously to (30). Since capital is perfectly mobile, the return to this factor will be equalized across all islands, therefore, we must have $r^I = r^A$ which implies:

$$\frac{k^I}{k^A} = \frac{\left(\mu + \frac{\beta}{2} - TAC^I(M^I) \right)}{Z \left(\mu + \frac{Z\beta}{2} \right)} \quad (37)$$

First, observe that if there were no agency costs in industry ($\sigma_I^2 = 0$), then $\frac{k^I}{k^A}$ and $\frac{M^I}{M^A}$ would be constant irrespective of the stock of capital of the economy. This implies that the perfect information version of this model would have sector-balanced growth. Next, consider the case with $\sigma_I^2 > 0$. In this case, as capital accumulates M^I grows, and from (37), $\frac{k^I}{k^A}$ increases. Therefore, there is faster capital deepening in industry than in agriculture. As a result, the shares of industrial production over total production and the share of expenditure in industrial goods over total expenditure also grow with development. Additionally, productivity and wages also increase in industry but remain constant in agriculture. As a result, economic growth in the presence of imperfect information is *endogenously sector-biased*, despite the fact that technical progress is neutral across the two sectors. This is consistent a very salient pattern in the development process of almost all countries: at the early stages of development, a large fraction of resources are allocated to agriculture, and as the economy grows, more resources are transferred to industry. This pattern of development is usually explained by assuming that the potential for productivity growth is much higher in manufacturing than in agriculture due

to some “sectorial externalities” (e.g. Matsuyama, 1991). Our mechanism can therefore be viewed as suggesting a microfoundation for these externalities.

4.3 From Villages to Cities

In the model discussed in the previous subsection, by construction, the share of total employment in agriculture remained constant. This feature is easy to change, and the model has interesting implications about migration from “rural villages” to “industrial cities”, another salient pattern of economic development. To do this, we introduce an additional factor of production, say *land*, which is immobile. In contrast to previous sections, we now assume that labor can move freely between islands, and agents can choose in which island to work. The interpretation of the model can also be modified to fit this example. Agricultural islands have low agency costs not only because of the different structure of uncertainty (as in the previous subsection), but also because of the way villages are organized: the close-knit communities lead to tight peer group monitoring. In contrast, the privacy and anonymity in cities do not allow easy direct monitoring, hence principal-agent relations have to use decentralized information and incentive contracts to induce effort. This point is also emphasized by Banerjee and Newman (1996) who obtain migration from villages to cities as a result of borrowing constraints becoming less severe.

To model these issues, let us modify the Leontieff part of the technology in both agricultural and industrial have to: $\min [1, q^{\alpha_q} l^{\alpha_l} k^{(1-\alpha_q-\alpha_l)}]$ instead of $\min [1, l^{\alpha} k^{(1-\alpha)}]$ (where q is land). In this case, both the rate of return to capital and wages will be equalized across islands, but the rental rate on land will not be.

The formal argument is very similar to that used to analyze the two factor case. Now, with labor perfectly mobile, wages and the rate of return on capital will be equalized across islands, and this implies:

$$w = \frac{\alpha_l Z \left(\mu + \frac{Z\beta}{2} \right)}{l^A} = \frac{\alpha_l \left(\mu + \frac{\beta}{2} - TAC^I(M^I) \right)}{l^I} \quad (38)$$

$$r = \frac{\alpha_k Z \left(\mu + \frac{Z\beta}{2} \right)}{k^A} = \frac{\alpha_k \left(\mu + \frac{\beta}{2} - TAC^I(M^I) \right)}{k^I} \quad (39)$$

where l^A (k^A) and l^I (k^I) denote labor (capital) employed by agricultural and

industrial firms, respectively. From (38), it follows that:

$$\frac{l^I}{l^A} = \frac{k^I}{k^A} = \frac{\left(\mu + \frac{\beta}{2} - TAC^I(M^I)\right)}{Z\left(\mu + \frac{Z\beta}{2}\right)} \quad (40)$$

Capital accumulation implies an increasing number of firms and decreasing agency costs in industry. Then, from (40), $\frac{l^I}{l^A}$ and $\frac{k^I}{k^A}$ have to grow. This implies that the number of firms in industry relative to agriculture, $\frac{M^I}{M^A}$, is growing. Therefore, total employment in agriculture is declining and agents must be migrating from villages to cities. The reason for this transformation is again related to our main mechanism: villages provide direct peer-monitoring thus reduce agency costs (i.e. without informational problems, there would be no migration). At the early stages of development, this is a very important resource and a large fraction of the population lives in villages, even though productivity is lower there (partly endogenously because there is excessive capital and labor in these islands). As more capital and information are accumulated, cities take advantage of reduced agency costs and expand. Finally, note that during the process of industrial expansion and migration, the productivity of capital and labor in agriculture is also increasing because fewer workers and capital are working with the same amount of land.

4.4 Direct Monitoring and Financial Development

At all stages of development, specialized financial institutions intermediate funds from savers to firms. However, there are very important differences between the institutions in poor economy and those in a more developed society. In his historical study Goldsmith (1987) finds that in most pre-modern societies funds are provided by direct lending institutions such as usurers, or local intermediaries, or at best, local banks. This contrasts with the larger banks, stock and bond markets of more developed economies. A crucial difference concerns the degree of direct monitoring carried out by different financial institutions (see Diamond, 1984, on the monitoring role of financial intermediaries).

In this section, we assume that there are two types of *financial* intermediaries, with free entry into each type. First, in each island there exist local credit institutions which we will call *village intermediaries* (VI). These intermediaries collect funds from savers in the whole economy at some market rate

r , but can only lend to firms located in their own island.⁹ A *VI* can perfectly monitor the effort of the local managers that they finance. The cost of providing these intermediation/monitoring services is c per firm. For simplicity, we assume that monitoring takes place *interim* and is publicly observed. This implies that a *VI* monitors all the managers it lends to before their final performance is revealed (more realistic assumptions, for instance, stochastic monitoring would not change the results). Second, there are some *global intermediaries* (*GI*) which offer their service at some lower cost, but cannot monitor (for instance, they lack the local expertise). For simplicity, and without loss of any generality, we assume that the cost of providing these services is zero. We can think of these *GIs* as of banks which operate at the economy-wide scale, or, alternatively, we can think that firms can raise their funds through a stock market (see Acemoglu and Zilibotti, 1996, for a more detailed analysis of these different institutions in a model with moral hazard problems). The important assumption here is that, in the absence of informational imperfections, intermediation through *GI* is more efficient than intermediation through *VI*. We also assume that the performance of firms which receive funds via *VIs* are observed by firms run through *GI*, and this assumption rules out potential coordination problem.

Clearly, the comparative advantage of local intermediation and monitoring declines as the scale of economic activity expands and more information is revealed by the activity of firms in the economy. We can prove the following proposition:

Proposition 7 *Let \hat{K} be such that $TAC(\hat{K}) = c$. Then, for all $K_t \leq \hat{K}$ intermediation is provided by “village intermediaries”, while for all $K_t \geq \hat{K}$ intermediation is provided by “global intermediaries”.*

This result is consistent with the empirical evidence that local financial institutions are predominant in poor economies and decline as development proceeds (see Besley, 1995, Fry, 1995, Goldsmith, 1987). Intuitively, *VIs* do not need decentralized information since they carry out direct monitoring.

⁹Many intermediaries may have local “expertise” which would justify this. The assumption that *VIs* borrow from savers in the whole economy may not be very realistic. In practice, they can do so by borrowing from other financial institutions. With *VIs* only using funds from their own islands, our analysis would be more involved, but the main result would not be affected.

Therefore, as more information is accumulated, global intermediaries become relatively more attractive. Note also that even though our simple model predicts an abrupt switch from local to global intermediation, the analysis could be easily extended to yield a smoother transition.

4.5 Division of Labor

Division of labor is a complex phenomenon, and different approaches concentrate on different aspects of this process. An important aspect of the division of labor is the delegation of tasks to agents who are not residual claimants of the returns they generate. In our economy, the most important form of delegation is to have managers running firms. In this extension we show that in poor economies where information is scarce, production techniques that do not rely on principal-agent relationships will have a comparative advantage and the extent of division of labor will be limited. This view is in line with a simple reading of historical patterns. During the First Industrial Revolution firms were predominantly family-managed, and this managerial structure is still dominant in many developing economies. More recently, especially starting with the Second Industrial Revolution, complex hierarchical organizations have emerged and played an important role in production and distribution (see Pollard, 1965, Chandler, 1977).

We analyze the issue of division of labor with a very simple extension model. We refer to our benchmark technology of production as the *factory production* (*FP*). The alternative is *primitive production* (*PP*) which entails no delegation of tasks to a manager. Output of unit i in island j using *PP* is given as:

$$y_{ijt} = (\mu + \theta + a_{jt}) \min \left[1, k_{ijt}^{1-\alpha} l_{ijt}^\alpha \right] + \epsilon_{ijt} \quad (41)$$

Since with *PP* there is no “fixed cost” of production, the number of firms using *PP* is indeterminate. However, to facilitate the discussion and without loss of generality, we impose that for all i, j , $l_{ijt}^\alpha k_{ijt}^{1-\alpha} = 1$, so that we can still talk of the “number of firms”.

We now assume that $\theta \in (\frac{\beta}{2} - TAC_\infty, \frac{\beta}{2} - TAC_1)$ where TAC_m is the total agency cost incurred with factory production when there are m firms in the island. This assumption ensures that when there are very few firms, *PP* is preferred to *FP*. In contrast, when there are sufficiently many firms so that

the society can set accurate standards, agency costs are low and division of labor (FP) is preferred to PP . This is because, when there are m active firms in the island, each firm with FP generates a certainty equivalent of income equal to $\mu + \frac{\beta}{2} - TAC_m$ whereas with PP , each firm produces a certainty equivalent of income equal to $\mu + \theta$. It is now clear that we can state an analogue of Proposition 7 whereby for all capital stocks less than some critical level \tilde{K} , the economy does not make use of division of labor, and when $K > \tilde{K}$, all production takes place with FP . \tilde{K} in this case is given by $TAC(\tilde{K}) = \frac{\beta}{2} - \theta$. Note also that at the point when the economy switches from PP to division of labor, there is an increase in the level of productivity per firm from $\mu + \theta$ to $\mu + e^*(\tilde{K})$ (where $e^*(\tilde{K}) = \frac{e^*(\tilde{K})^2}{2\beta} - \frac{\rho}{2}Var(\zeta^*(\tilde{K})) = \theta$).

It is useful to emphasize that in our economy the switch from primitive production to division of labor is *not* due to the fact that division of labor is more capital intensive but that it is more *information intensive*. In other words, in the absence of informational imperfections, if PP was preferred to FP at some capital level K_0 , it would be preferred at all other capital levels too. Hence, loosely speaking, the division of labor in our economy is limited by the extent of information: when the economy has more capital, and thus more information, agency costs decline and division of labor becomes relatively more attractive.

The analysis of the last two subsections also suggests a more general principal. If more “sophisticated” products or production techniques are at the same time more “information intensive”, because monitoring is harder or because they involve more delegation, then as a society develops, the range of products will expand, the production methods will become more refined and there will be more delegation of tasks. With all of these changes, as in our division of labor example, the productivity will often increase.

4.6 Development and Specialization

Less developed economies are typically highly specialized and invest a large share of their resources in only a few narrow sectors. This is consistent with our model when we view islands as sectors. In particular, the analysis of section 3 showed that asymmetric equilibria are possible with imperfect information, even though with perfect information, only symmetric equilibria can exist.

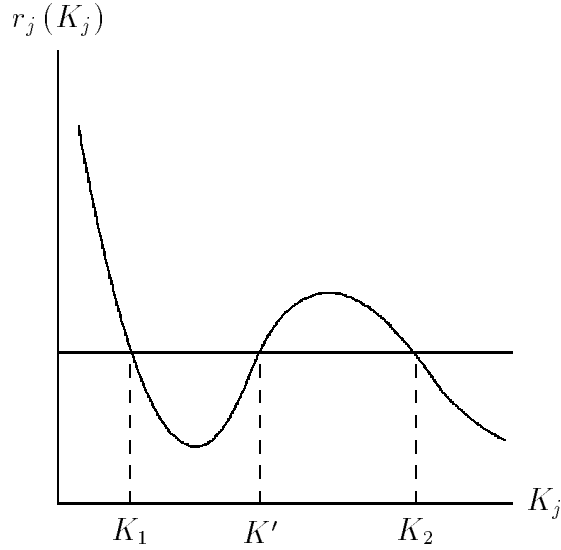


Figure 1: Asymmetric equilibria.

Asymmetric equilibria correspond to the economy being specialized in a few activities at the expense of the rest. The reason is to economize on agency costs by generating information in some selected sectors.

To get more insights into development and specialization, we can consider a simple diagrammatic exposition of asymmetric equilibria. Figure 1 draws $r_j(K_j)$, the rate of return to capital in island j if an amount of capital K_j is invested in that island (thus with $M(K_j)$ active firms). It is straightforward to see that if this schedule is downward sloping everywhere, there can be no asymmetric equilibria, and our condition in Lemma 4, that $\mu > \bar{\mu}$, ensures this. Suppose instead that $\mu < \bar{\mu}$ and $r_j(K_j)$ is increasing over a certain range. Let us now concentrate on asymmetric equilibrium in which a fraction λ of islands have capital K_2 and the remaining $1 - \lambda$ have $K_1 < K_2$, and $K \in (K_1, K_2)$ so that not all islands can have capital equal to K_2 . It is straightforward to see that K_1 and K_2 must be as drawn in Figure 1. In particular, note that no island will have a capital stock equal to K' because $r_j(K_j)$ is increasing around K' , thus if an additional firm opened in this island, it would make positive profits. Then, an asymmetric equilibrium is characterized as follows: fix K_1 and K_2 as in Figure 1, then compute:¹⁰

¹⁰As the argument suggests, when one asymmetric equilibrium exists, there will also exist many others.

$$\lambda(K) = \frac{K - K_1}{K_2 - K_1}$$

That is the fraction of islands with more capital is determined in order to have the capital demand equal the aggregate stock of capital. It is now clear that conditional on K_1 and K_2 , a smaller aggregate stock of capital implies that the economy is very specialized. For instance, thinking of the islands as sectors, this would correspond to an economy which invests a large fraction of its resources in a small fraction of the sectors. As the economy becomes richer and K increases, $\lambda(K)$ increases too, and the economy becomes less specialized. Therefore, our model predicts that poorer societies tend to be more specialized as a way of economizing on agency costs.

5 Conclusion

This paper offers a theory of development where principal-agent relations play a crucial role. Wealth is generated by delegating tasks to agents who are not the residual claimants of the returns they generate. When the control of these agents is costly, productivity is low. We argue that the amount of decentralized information the society generates is a crucial determinant of how easy it is to control the agents. In turn, the structure of information depends on the scale of production. When more agents are engaged in the same activity, the society will develop a better *standard* for this activity, and this will enable relative performance evaluation. Therefore, as a society accumulates more capital, it will also accumulate more information. This will lead to higher managerial effort and productivity.

As well as explain why the process of development may be slow, our theory has a number of important implications, reminding us of the older theories of development with their emphasis on structural transformation. We find that the extent of risk sharing will change with development, the sectorial composition of output will shift, there will be more division of labor, there will be less reason for close-knit village communities to survive, and financial institutions will be transformed. Approaching some of the same problems as our paper with a somewhat different emphasis, Besley (1995, p.121) writes that “(local institutions and enforcement)... *do seem in general do disappear*

as capital markets develop. This reflects the fact that monitoring and other technologies improve in the development process. ... Whether a symptom or a cause, the decline of this type of non-market institution in the development process vividly illustrates the idea that they use certain information structures and enforcement technologies that are eroded by the transformation to a modern economy." In terms of Besley's quotation, we are arguing that the relative decline of a host of institutions and sectors is a *consequence* of information accumulation and development, but also that such structural transformations have important implications regarding the range of products, organization of firms and productivity.

Our list is not even close to being exhaustive of the transformations on the way from poverty to prosperity. It is also not exhaustive of the development implications of changing principal-agent relations. Our model is sufficiently simple and tractable that more results can be obtained by modifying certain aspects of the baseline specification. However, we hope the implications we draw give the flavor of what can be done. We also hope that our model suggests other approaches to the same problem: more information will improve agency relations not only via better relative performance evaluation but also through alternative uses of information. Some of the implications will be similar while others will differ. As different testable implications are derived from these approaches and are confronted with data, our understanding of the development process will improve.

Appendix

Proof of Lemma 1. *Part 1.* Since markets are complete, firms and managers will agree on the first-best level of effort which, from (2) and (3), is $e_{ijt} = \beta$.

Part 2. We start by showing that k_j , the capital per firm in island j , is a monotonically increasing function of M_j , the number of firms in this island. Given this, we prove that all islands have the same number of firms using proof by contradiction.

First, since firms compete to maximize profits, we must have in the second-stage game that $\frac{w_j}{r} = \frac{\partial y_{ij}/\partial l_{ij}^-}{\partial y_{ij}/\partial k_{ij}^-}$ (where the superscript $-$ indicates that these are partial derivatives “from below”). Then, from the unitary elasticity of substitution between labor and capital, $\frac{k_{ij}}{l_{ij}} = \frac{k_j}{l_j} = \frac{1-\alpha}{\alpha} \frac{w_j}{r}$ (all firms in island j adopt the same technology). Since when (6) holds all firms are productively efficient ($l_j^\alpha k_j^{1-\alpha} = 1$), it must be the case from (4) that $l_{ij} = l_j = \frac{N}{M_j}$ and $k_{ij} = k_j = \frac{K_j}{M_j}$. Then, aggregating within each island, we have that:

$$N^\alpha K_j^{1-\alpha} = M_j \quad (42)$$

which can be rearranged to give:

$$k_j = \left(\frac{K_j}{N}\right)^\alpha \text{ and } l_j = \left(\frac{N}{K_j}\right)^{1-\alpha} \quad (43)$$

Now, suppose that there are two islands, j', j'' such that $K_{j'} > K_{j''}$. (42) implies that $M_{j'} > M_{j''}$. This implies from (43) that $l_{j'} < l_{j''}$ and $k_{j'} > k_{j''}$. Therefore, $\frac{k_{j'}}{l_{j'}} > \frac{k_{j''}}{l_{j''}}$ and because in both islands managers exert $e = \beta$, the rate of return to capital in firms of island j' is lower than in j'' . Since the rate of return to capital has to be equalized across islands, this gives a contradiction. Therefore, we must have $K_{j'} = K_{j''} = K_t$ and $M_{j'} = M_{j''} = M_t$. Hence, $l_j = \left(\frac{N}{K_t}\right)^{1-\alpha}$ and $k_j = \left(\frac{K_t}{N}\right)^\alpha$.

Part 3. From part 2, $\frac{rk_j}{w_j l_j} = \frac{1-\alpha}{\alpha}$. Free-entry in the first stage game implies that all firms must be making zero expected profits. Therefore, $rk_j = (1-\alpha)E[y_j - (z_j - w_j)]$ and $w_j l_j = \alpha E[y_j - (z_j - w_j)]$. Since, agents must be indifferent between becoming managers or workers, hence $z_j = w_j + \frac{e_j^2}{2\beta}$.

Now using the facts that $y_j = \mu + e_j$, $e_j = \beta$ and that the equilibrium must be symmetric, we obtain (8), (9) and (10). This concludes the proof of Lemma 1. QED

Proof of Proposition 1: By assumption $K_0 \geq N^{\frac{-\alpha}{1-\alpha}}$. Then condition (6) ensures that if K_t is in the right neighborhood of $N^{\frac{-\alpha}{1-\alpha}}$, $K_{t+1} > K_t$. Next, given (7), K_{t+1} is an increasing and strictly concave function of K_t and since $M_t \leq N$, we have $K_t \leq s(\mu + \beta)N$. Therefore, there exists a unique steady state level of K_t , K^{ss} . Since K_{t+1} is a strictly concave function, this unique steady state is also globally stable.

To characterize the steady state value K^{ss} , note that $K^{ss} = s(\mu + \beta)M(K^{ss})$. Then using (7) gives the expression of K^{ss} in the proposition. The rest of the Proposition follows immediately from the analysis discussed in the text. QED.

Proof of Lemma 2: *Part 1.* The proof is identical to the proof of Lemma 1 part 2, except for the symmetry argument.

Part 2. Since firms are risk-neutral, by a standard argument, they will offer a non-random return to the factors of production for which there are no incentive compatibility constraints. Therefore, the rates of return to labor and capital are non-random. The exact expressions for these rates of return, (15) and (16), follow by the same argument as part 3 of Lemma 1. QED.

Proof of Lemma 3. *Part 1.* The utility of manager i in island j is given by $U_i(e_{ij}, \cdot) = E(\zeta_{ij} | e_{ij}) + w_j - \frac{1}{2\beta}e_{ij}^2 - \frac{\rho}{2}Var(\zeta_{ij}) = \Theta + (\phi_{1ij} + \phi_{2ij})(\mu + e_{ij}) - \frac{1}{2\beta}e_{ij}^2$, where ζ_{ij} is given by (17), and Θ collects terms that do not depend on e_{ij} . Since the manager chooses effort to maximize U_i , we have $e_{ij}^* = \beta(\phi_{1ij} + \phi_{2ij})$. *Part 2* is straightforward. *Part 3.* $Var(\zeta_{ij}) = E[\zeta_{ij} - \phi_{0ij}]^2 = E[\phi_{1ij}(a_j + \epsilon_{ij}) + \phi_{2ij}(\epsilon_{ij} - \sum_{s \neq i} \epsilon_{sj})]^2 = [(\phi_{1ij} + \phi_{2ij})^2 \sigma^2 + \phi_{1ij}^2 \nu^2 + \phi_{2ij}^2 \frac{\sigma^2}{M_j - 1}]$. QED.

Proof of Proposition 2. Using Lemma 3, we write the maximization of (21) as:

$$\max_{\phi_{1ij}, \phi_{2ij}} [\mu + \beta(\phi_{1ij} + \phi_{2ij})] - \frac{\beta}{2}(\phi_{1ij} + \phi_{2ij})^2 - \frac{\rho}{2} \left[(\phi_{1ij} + \phi_{2ij})^2 \sigma^2 + \phi_{1ij}^2 \nu^2 + \phi_{2ij}^2 \frac{\sigma^2}{M_j - 1} \right] \quad (44)$$

Solving the two first-order conditions gives $\phi_{1ij} = \phi_{1j}^*$ and $\phi_{2ij} = \phi_{2j}^*$ as in (24) and (25). To find $\phi_{0ij} = \phi_{0j}^*$, we use the participation constraint, (20), and the facts that $e_{ij}^* = \beta(\phi_{1j}^* + \phi_{2j}^*)$ (from Lemma 3) and $E(\zeta \mid e_{ij}^*) = \phi_{0j}^*$ (from (17)). QED.

Proof of Lemma 4. Let us define the rate of return to capital in island j when a total amount of capital K_j is invested there, the labor market clears, firms choose the optimal contracts and make zero profits as:

$$\begin{aligned} r_j(K_j) &= (1 - \alpha)(\mu + \beta(\phi_1^*(K_j) + \phi_2^*(K_j)) - \phi_0^*(K_j)) \left(\frac{N}{K_j}\right)^\alpha \\ &= (1 - \alpha)\mu \left(\frac{N}{K_j}\right)^\alpha + (1 - \alpha)\beta[(\phi_1^*(K_j) + \phi_2^*(K_j)) - \phi_0^*(K_j)] \left(\frac{N}{K_j}\right)^\alpha \end{aligned} \quad (45)$$

The fact that capital should receive the same rate of return in all islands implies that for all $j \in [0, 1]$: $r_j(K_j) = r(\mathbf{M}_t, K_t)$.

A necessary condition for asymmetric equilibrium is that there exist two levels of capital, $K_{j'}$, $K_{j''}$, such that $r_{j'}(K_{j'}) = r_{j''}(K_{j''})$. Therefore, a sufficient condition for the equilibrium to be unique and symmetric is that $r'_j(K_j) < 0$ for all K_j . We will now prove that for μ sufficiently large this is always the case. To see this note that: (i) the first term on the RHS of (45), $(1 - \alpha)\mu \left(\frac{N}{K_j}\right)^\alpha$, is decreasing in K_j ; (ii) the second term of (45), $\Psi(K_j) \equiv (1 - \alpha)\beta[(\phi_1^*(K_j) + \phi_2^*(K_j)) - \phi_0^*(K_j)] \left(\frac{N}{K_j}\right)^\alpha$, does not depend on μ ; (iii) from (24), (25) and (14) it follows that $\exists B^u$ such that for any $K_j > 0$, $\frac{d\Psi(K_j)}{dK_j} < B^u < \infty$. Then, $\exists \bar{\mu}$ such that $\forall \mu \geq \bar{\mu}$ we have that $r'_j(K_j) < 0$, and there exists a unique equilibrium whereby $\forall j \in [0, 1]$, $K_j = K$ (capital is equally distributed across islands). QED.

Proof of Proposition 4. Recall that ϕ_1 and ϕ_2 are chosen to maximize (21) subject to (19). Given the definition of TAC from (30), this is equivalent to the unconstrained maximization of $\mu + \frac{\beta}{2} - TAC(K)$ with respect to ϕ_1 and ϕ_2 . Now recall from Lemma 3 that $e^* = \beta(\phi_1^* + \phi_2^*)$ and $Var(\zeta^*) = \left[((\phi_1^* + \phi_2^*)^2 \sigma^2 + (\phi_1^*)^2 \nu^2 + (\phi_2^*)^2 \frac{\sigma^2}{M-1}) \right]$. Then, we can write $TAC(K) = TAC(M(K), \phi_1^*(K), \phi_2^*(K))$. The envelope theorem implies that $\frac{\partial TAC}{\partial \phi_1^*} \frac{d\phi_1^*}{dK} = \frac{\partial TAC}{\partial \phi_2^*} \frac{d\phi_2^*}{dK} = 0$. Therefore $TAC'(K) = \frac{\partial TAC}{\partial M(K)} M'(K) = -\frac{\rho}{2} (\phi_2^*)^2 \frac{\sigma^2}{(M-1)^2} (1 - \alpha) \left(\frac{N}{K}\right)^\alpha < 0$.

For the second part, from the definition of SAC we have that:

$$SAC(\bar{e}, K) = \min_{\phi_2} \left[(\sigma^2 + \nu^2) \left(\frac{\bar{e}}{\beta} \right)^2 + \left(\frac{\bar{e}}{\beta} - \phi_2 \right)^2 \nu^2 + \frac{\sigma^2}{M(K) - 1} \phi_2^2 \right]$$

Differentiating this with respect to K and once more using the envelope theorem, we have $SAC_2(\bar{e}, K) < 0$. QED.

Proof of Proposition 5. Conditional upon \mathbf{M}_t , ϕ_{1j} and ϕ_{2j} are given by the first-order conditions of (32) once e_j is substituted from (19). Straight-forward differentiation leads to (24) and (25) exactly as in the decentralized equilibrium. QED

Proof of Proposition 6. From (23), (24) and (25), it follows that:

$$V(K) = \sigma^2 \frac{(\sigma^2 + (M - 1)\nu^2)(\sigma^2 + M\nu^2)}{\left[(\sigma^2 + (M - 1)\nu^2) + (\sigma^2 + M\nu^2) \frac{\rho\sigma^2}{\beta} \right]^2}.$$

Let Υ_N denote the numerator and Υ_D the denominator of the right hand-side expression. Then:

$$\begin{aligned} V'(K) &= \frac{\sigma^2\nu^2}{\Upsilon_D^3} \left\{ [(\sigma^2 + (M - 1)\nu^2) + (\sigma^2 + M\nu^2)] \Upsilon_D - 2 \left(1 + \frac{\rho\sigma^2}{\beta} \right) \Upsilon_N \right\} \\ &= \frac{\sigma^2\nu^4}{\Upsilon_D^3} \left[(\sigma^2 + M\nu^2) \frac{\rho\sigma^2}{\beta} - (\sigma^2 + (M - 1)\nu^2) \right] \end{aligned}$$

where some straightforward algebra is necessary to go from the first to the second line. This expression establishes that, $\text{sign}[V'(K)] = \text{sign} \left[\nu^2 - \left(1 - \frac{\rho\sigma^2}{\beta} \right) (\sigma^2 + M\nu^2) \right]$. If $\rho\sigma^2 > \beta$, then $V'(K) > 0$ for all K which proves part 3. Next, set $M = 1$. If $\rho(\sigma^2 + \nu^2) < \beta$, then $V'(K \mid M(K) = 1) < 0$, and $V'(K) < 0$ for all K such that $M(K) \geq 1$ (part 1). Finally, if $\rho\sigma^2 < \beta < \rho(\sigma^2 + \nu^2)$, we have $V'(K \mid M(K) = 1) > 0$, but since $\frac{\rho\sigma^2}{\beta} < 1$, $V'(K)$ is decreasing in M (thus K). Moreover, for M (or equivalently K) sufficiently large $V'(K) < 0$. Therefore, there exists \bar{K} such that for $K < \bar{K}$, $V'(K) > 0$ and for $K > \bar{K}$, $V'(K) < 0$ (part 2). QED.

Proof of Proposition 7. First, consider $K_t < \hat{K}$ as defined in the proposition. We will show that all intermediation is through VIs is the unique equilibrium. Suppose all intermediation through VIs . In this case, free entry ensures that active firms which are using VIs make zero profit, i.e. $\left(\mu + \frac{\beta}{2} - c \right) =$

$w^{VI}(K_t)l(K_t) + r^{VI}k(K_t)$, where w^{VI} and r^{VI} are the equilibrium factor returns when there is only intermediation through VI . Now, consider a deviation from a firm which decides to use GI instead of VI . The profit of this firm will be $\left(\mu + \frac{\beta}{2} - TAC(K_t)\right) - w^{VI}(K_t)l(K_t) - r^{VI}k(K_t) = c - TAC(K_t) < 0$, since $TAC(K_t) > TAC(\hat{K}) = c$. Hence, intermediation through VI s is an equilibrium. We then show that in the same case ($K_t \leq \hat{K}$) all intermediation through GI s is not an equilibrium. Assume it is, then free entry in the first-stage game ensures that active firms which are using GI s make zero profit, i.e. $\left(\mu + \frac{\beta}{2} - TAC(K_t)\right) = w^{GI}(K_t)l(K_t) - r^{GI}k(K_t)$. Now, consider a deviation from a firm which decides to use VI instead of GI . The profit of this firm will be $\left(\mu + \frac{\beta}{2} - c\right) - w^{GI}(K_t)l(K_t) - r^{GI}k(K_t) = TAC(K_t) - c > 0$. This establishes that there exists a profitable deviation, therefore intermediation through GI s is not an equilibrium. A similar argument would show that no equilibrium in which some firms use GI s and some others use VI s can exist. Thus, with $K_t < \hat{K}$, intermediation through VI s is the unique equilibrium.

Next consider $K_t > \hat{K}$. In this case, the reverse of the previous argument applies exactly, and this establishes that only intermediation through GI s is an equilibrium. QED

References

- [1] Acemoglu, D., and F. Zilibotti (1996). ‘Agency Costs in the Process of Development.’ CEPR Discussion Paper No. 1421.
- [2] Acemoglu, D., and F. Zilibotti (1997). ‘Was Prometheus Unbound by Chance? Risk, Diversification and Growth.’ forthcoming *Journal of Political Economy*, August 1997.
- [3] Andreoni, J. (1989). ‘Giving with Impure Altruism: Application to Charity and Ricardian Equivalence.’ *Journal of Political Economy*, 97, 1447-1458.
- [4] Backus, D., P. Kehoe and T. Kehoe (1992). ‘In Search of Scale Effects in Trade and Growth’ *Journal of Economic Theory*, 58, 377-409.
- [5] Bairoch, P. (1988). *Cities and Economic Development: From the Dawn of History to the Present*, translated by C. Braiden, Chicago University Press: Chicago.
- [6] Banerjee, A. and A. Newman (1993). ‘Occupational Choice and the Process of Development.’ *Journal of Political Economy*, 101, 274-298.
- [7] Banerjee, A. and A. Newman (1996). ‘A Dual Economy Model of Modernization and Development.’ Mimeo, MIT.
- [8] Besley, T. (1995). ‘Non-Market Institutions for Credit and Risk Sharing in Low Income Countries’ *Journal of Economic Perspectives*, 9, 115-128.
- [9] Chandler, A. (1977). *The Visible Hand: The Managerial Revolution*. Oxford University Press, Oxford.
- [10] Diamond, D. (1984). ‘Financial Intermediation and Delegated Monitoring.’ *Review of Economic Studies*, 51, 393-414.
- [11] Fry, M. (1995). *Money, Interest, and Banking in Economic Development*. Second Edition. The Johns Hopkins University Press, Baltimore.

- [12] Green, J. and N. Stokey (1983). ‘A Comparison of Tournaments and Contests.’ *Journal of Political Economy*, 91, 349-64.
- [13] Greenwood, J. and B. Jovanovic (1990). “Financial Development, Growth and the Distribution of Income.” *Journal of Political Economy* 98, 1067-1107.
- [14] Greenwood, J. and B. Smith (1993). “Financial Markets in Development and the Development of Financial Markets.” Mimeo University of Rochester and Cornell.
- [15] Goldsmith, R. (1987). *Premodern Financial Systems: A Historical Comparative Study*, Cambridge University Press, Cambridge.
- [16] Holmstrom, B. (1979). ‘Moral Hazard and Observability.’ *Bell Journal of Economics* , 10, 74-91.
- [17] Holmstrom, B. (1982). ‘Moral Hazard in Teams.’ *Bell Journal of Economics*, 13, 323-340.
- [18] Holmstrom, B. and P. Milgrom (1987). ‘Aggregation and Linearity in the Provision of Intertemporal Incentives.’ *Econometrica*, 55, 303-28.
- [19] Holmstrom, B. and P. Milgrom (1991). ‘Multitask Principal-Agent Analyses: Incentive Contracts, Asset Ownership and Job Design.’ *Journal of Law and Economics*, 7 (special).
- [20] Lazaer, E. and S. Rosen (1981). ‘Rank-Order Tournaments as Optimal Labor Contracts.’ *Journal of Political Economy*, 89, 841-64.
- [21] Matsuyama, K. (1991). ‘Increasing Returns, Industrialization and Indeterminacy of Equilibrium.’ *Quarterly Journal of Economics*, 106, 617-650.
- [22] Matsuyama, K. (1995). ‘Complementarities and Cumulative Processes in Models of Monopolistic Competition.’ *Journal of Economic Literature*, 33, 701-729.
- [23] Mokyr, J. (1991). *British Industrial Revolution: An Economic Perspective*. Westview, Boulder.

- [24] Morduch, J. (1995). 'Income Smoothing and Consumption Smoothing.' *Journal of Economic Perspectives*, 9, 103-114.
- [25] Murphy, K. M., A. Shleifer and R. Vishny (1989). 'Industrialization and the Big Push' *Journal of Political Economy*, 97, 1003-1026.
- [26] Myrdal, G. (1968). *The Asian Drama: An Inquiry into the Poverty of Nations*. The 20th Century Fund, New York.
- [27] North, D. (1990). *Institutions, Institutional Change and Economic Performance*, Cambridge University Press, Cambridge.
- [28] Pollard S., (1965) *The Genesis of Modern Management: A Study of Industrial Revolution in Great Britain* Harvard University Press
- [29] Rosenberg, N., and L. Bridzell Jr. (1985). *How the West Grew Rich*. New York Basic Books.
- [30] Shleifer, A. (1985). 'A Theory of Yardstick Competition.' *Rand Journal of economics*, 16, 319-27.
- [31] Singer, H. (1949). 'Economic Progress in Underdeveloped Countries' *Social Research*.
- [32] Stiglitz, J. (1987). 'Economic Organization, Information and Development.' In Chenery, H. and T. Srinivasan, eds. *Handbook of Development Economics* Vol. I, 93-160. Amsterdam: North Holland.
- [33] Townsend, R. (1983). "Theories of Intermediated Structures" *Carnegie-Rochester Conference Series on Public Policy* 18, 221-272.
- [34] Townsend, R. (1994). 'Risk and Insurance in Village India.' *Econometrica*, 62, 539-592.
- [35] Townsend, R. (1995a). 'Financial Systems in Northern Thai Villages.' *Quarterly Journal of Economics*, 110, 1011-1046.
- [36] Townsend, R. (1995b). 'Consumption Insurance: An Evaluation of Risk-Bearing Systems in Low-Income Economies.' *Journal of Economic Perspectives*, 9, 83-102.